UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/345,238	06/30/1999	SCOTT SHAOBING CHEN	YO999-172	9988

. 7590 12/17/2003

KEVIN M MASON
RYAN, MASON & LEWIS, LLP
1300 POST ROAD, SUITE 205
FAIRFIELD, CT 06430

EXAMINER	
HAN, QI	

ART UNIT	PAPER NUMBER
2654	18

DATE MAILED: 12/17/2003

Please find below and/or attached an Office communication concerning this application or proceeding.

UNITED STATES PATENT AND TRADEMARK OFFICE

COMMISSIONER FOR PATENTS
UNITED STATES PATENT AND TRADEMARK OFFICE
P.O. BOX 1450
ALEXANDRIA, VA 22313-1450
www.uspto.gov

BEFORE THE BOARD OF PATENT APPEALS AND INTERFERENCES

Paper No. 18

Application Number: 09/345,238
Filing Date: June 30, 1999
Appellant(s): CHEN ET AL.

Kevin M. Mason
For Appellant

EXAMINER'S ANSWER

This is in response to the appeal brief filed on 10/6/2003.

(1) Real Party in Interest

A statement identifying the real party in interest is contained in the brief.

(2) Related Appeals and Interferences

The brief does not contain a statement identifying the related appeals and interferences which will directly affect or be directly affected by or have a bearing on the decision in the pending appeal is contained in the brief. Therefore, it is presumed that there are none. The Board, however, may exercise its discretion to require an explicit statement as to the existence of any related appeals and interferences.

(3) Status of Claims

The statement of the status of the claims contained in the brief is correct.

(4) Status of Amendments After Final

The appellant's statement of the status of amendments after final rejection contained in the brief is correct.

(5) Summary of Invention

The summary of invention contained in the brief is correct.

(6) Issues

The appellant's statement of the issues in the brief is correct.

(7) Grouping of Claims

The rejection of claims 1-35 stand or fall together because appellant's brief does not include a statement that this grouping of claims does not stand or fall together and reasons in support thereof. See 37 CFR 1.192(c)(7).

(8) Claims Appealed

The copy of the appealed claims contained in the Appendix to the brief is correct.

(9) Prior Art of Record

Chen et al. ("speaker, Environment and Channel Change Detection and Cluster via the Bayesian Information Criterion," proceedings of the DARPA broadcast news transcription and understanding workshop, Lansdowne, VA, Feb 8-11, 1998)

Kleider et al. (USPN 5,930,748, Filed : 07/11/1997)

(10) Grounds of Rejection

The following ground(s) of rejection are applicable to the appealed claims:

Claim Rejections - 35 USC § 102

The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

> A person shall be entitled to a patent unless –
>
> (b) the invention was patented or described in a printed publication in this or a foreign country or in public use or on sale in this country, more than one year prior to the date of application for patent in the United States.

1. Claims 1-5, 8, 10-14, 16-19, 21-26 and 28-35 are rejected under 35 U.S.C. 102(b) as being anticipated by Chen et al. ("speaker, Environment and Channel Change Detection and Cluster via the Bayesian Information Criterion," proceedings of the DARPA broadcast news transcription and understanding workshop, Lansdowne, VA, Feb 8-11, 1998) hereinafter referenced as Chen.

Regarding **claim 1**, Chen discloses speaker, Environment and Channel Change Detection and Cluster via the Bayesian Information Criterion for segmenting the audio stream into homogeneous region according to speaker identity, environmental condition and channel condition and clustering speech segments into homogeneous clusters according to speaker

identity, environmental condition and channel (page 1, paragraph 2), which is read on the claimed "a method of tracking a speaker in an audio source, said method comprising the steps of: identifying potential segment boundaries in said audio source; and clustering homogeneous segments from said audio source substantially concurrently with said identifying step."

Regarding **claim 2**, Chen discloses everything claimed, as applied above (see claim 1). Chen further discloses that decision for detecting changes in speaker identity is based on the Bayesian Information Criterion (BIC) (page 2, paragraph 4), which is read on the claimed" wherein said identifying step identifies segment boundaries using a BIC model-selection criterion."

Regarding **claim 3**, Chen discloses everything claimed, as applied above (see claim 2). Chen further assumes that the sequence of cepstral vectors is draw from an independent multivariate Gaussian process and there is at most one changing point in the Gaussian process (page 3, paragraphs 4-5), and discloses that the hypothesis testing is viewed as a problem of model selection by comparing two models: one models the data as two Gaussians; the other models the data as just one Gaussian (page 4, paragraph 2), which is read on the claimed "wherein a first model assumes there is no boundary in a portion of said audio source and a second model assumes there is a boundary in said portion of said audio source."

Regarding **claim 4**, Chen discloses everything claimed, as applied above (see claim 2). Chen further discloses a combination of two equations: the maximum likelihood ratio (page 3, equation (2)) and the difference between the BIC values of the two models (page 4, equation (3)), which is inherently equivalent to the equation as claimed.

Regarding **claim 5**, Chen discloses everything claimed, as applied above (see claim 1). Chen further discloses that an algorithm sequentially detect the changing points in the Gaussian process and suggests that the algorithm starts with a small window and then extends the window size in each detecting loop (page 6, paragraph 1). It is also inherently true that the smaller the window size, the more unlikely the segment boundary occurs. This is read on the claimed "identifying step considers a smaller window size, n, of samples in areas where a segment boundary is unlikely to occur."

Regarding **claim 8**, Chen discloses everything claimed, as applied above (see claim 2). Chen further suggests not using the detected change point in new process window (see the algorithm: set a= t + 1) (page 6, paragraph 1), which is read on the claimed "BIC model selection test is not performed at the border of each window of samples."

Regarding **claim 10**, Chen discloses everything claimed, as applied above (see claim 1). Chen further discloses to apply the BIC criterion in clustering (page 8, paragraph 2), which is read on the claimed "clustering step is performed using a BIC model-selection criterion."

Regarding **claim 11**, Chen discloses everything claimed, as applied above (see claim 10). Chen further discloses that in the hierarchical clustering two nodes can be merged only if the merging increases the BIC value (abstract, also see page 9, paragraph 3) that suggests the two models used in identifying step are also applied in clustering step, which is read on the claimed "wherein a first model assumes that two segments or clusters should be merged, and a second model assumes that said two segments or clusters should be maintained independently."

Regarding **claim 12**, Chen discloses everything claimed, as applied above (see claim 11). Chen further discloses that the two nodes should not merger if an equation (8) (page 9,

paragraph 4) is negative, which is read on the claimed "merging said two clusters if a difference in BIC values for each of said models is positive."

Regarding **claim 13**, Chen discloses everything claimed, as applied above (see claim 1). Chen further discloses that using M segments and k clusters (page 8, paragraphs 2 and 3) successively merge two nearest nodes in clustering step and generate a new cluster set S' from pervious set S (page 9, paragraph 3), which is read on the claimed "clustering step is performed using K previously identified clusters and M segments to be clustered."

Regarding **claim 14**, Chen discloses everything claimed, as applied above (see claim 1). Chen further suggests to assign s as an identifier for a new cluster from two previous nodes or clusters s1 and s2 after each merging (page 9, paragraph 3), which is read on the claimed "the step of assigning a cluster identifier to each of said clusters." In addition, it is inherently true that an index of data structure employed for clustering task can be always used as a cluster identifier in software and/or firmware based process.

Regarding **claim 16**, the rejection bases on the same reason as applied above (see claim 1) because Chen discloses the same method for both "segments from said audio source corresponding to the same speaker" and "homogeneous segments". In addition, the applicant points out that "humongous segments" are "generally corresponding to the same speaker" (abstract).

Regarding **claim 17**, Chen discloses everything claimed, as applied above (see claim 16). Chen further discloses that decision for detecting changes in speaker identity is based on the Bayesian Information Criterion (BIC) (page 2, paragraph 4), which is read on the claimed

"wherein said identifying step identifies segment boundaries using a BIC model-selection criterion."

Regarding **claim 18**, Chen discloses everything claimed, as applied above (see claim 17). Chen further assumes that the sequence of cepstral vectors is draw from an independent multivariate Gaussian process and there is at most one changing point in the Gaussian process (page 3, paragraphs 4-5), and discloses that the hypothesis testing is view as a problem of model selection from two models: one models the data as two Gaussians; the other models the data as just one Gaussian (page 4, paragraph 2), which is read on the claimed "wherein a first model assumes there is no boundary in a portion of said audio source and a second model assumes there is a boundary in said portion of said audio source."

Regarding **claim 19**, Chen discloses everything claimed, as applied above (see claim 16). Chen further discloses that an algorithm sequentially detect the changing points in the Gaussian process and suggests that the algorithm starts with a small window and then extends the window size in each detecting loop (page 6, paragraph 1). It is also inherently true that the smaller the window size, the more unlikely the segment boundary occurs. This is read on the claimed "identifying step considers a smaller window size, n, of samples in areas where a segment boundary is unlikely to occur."

Regarding **claim 21**, Chen discloses everything claimed, as applied above (see claim 16). Chen further discloses to apply the BIC criterion in clustering (page 8, paragraph 2). Moreover, Chen discloses that in the hierarchical clustering two nodes can be merged only if the merging increases the BIC value (abstract, also see page 9, paragraph 3) that suggests the two models used in identifying step are also applied in clustering step, which is read on the claimed

"clustering step is performed using a BIC model-selection criterion, where a first model assumes that two segments or clusters should be merged, and a second model assumes that said two segments or clusters should be maintained independently."

Regarding **claim 22**, Chen discloses everything claimed, as applied above (see claim 16). Chen further discloses that using M segments and k clusters (page 8, paragraphs 2 and 3) successively merge two nearest nodes in clustering step and generate a new cluster set S' from pervious set S (page 9, paragraph 3), which is read on the claimed "clustering step is performed using K previously identified clusters and M segments to be clustered."

Regarding **claim 23**, the rejection bases on the same reason as applied above (see claim 16) because the same method in Chen's disclosure can also be applied for claim 23 "the steps of: identifying potential segment boundaries during a pass through said audio source; and clustering segments from said audio source corresponding to the same speaker during said pass through said audio source."

Regarding **claim 24**, Chen discloses everything claimed, as applied above (see claim 23). Chen further discloses that decision for detecting changes in speaker identity is based on the Bayesian Information Criterion (BIC) (page 2, paragraph 4), which is read on the claimed "said identifying step identifies segment boundaries using a BIC model-selection criterion."

Regarding **claim 25**, Chen discloses everything claimed, as applied above (see claim 24). Chen further assumes that the sequence of cepstral vectors is draw from an independent multivariate Gaussian process and there is at most one changing point in the Gaussian process (page 3, paragraphs 4-5), and discloses that the hypothesis testing is view as a problem of model selection from two models: one models the data as two Gaussians; the other models the data as

just one Gaussian (page 4, paragraph 2), which is read on the claimed "wherein a first model assumes there is no boundary in a portion of said audio source and a second model assumes there is a boundary in said portion of said audio source."

Regarding **claim 26**, Chen discloses everything claimed, as applied above (see claim 23). Chen further discloses that an algorithm sequentially detect the changing points in the Gaussian process and suggests that the algorithm starts with a small window and then extends the window size in each detecting loop (page 6, paragraph 1). It is also inherently true that the smaller the window size, the more unlikely the segment boundary occurs. This is read on the claimed "identifying step considers a smaller window size, n, of samples in areas where a segment boundary is unlikely to occur."

Regarding **claim 28**, Chen discloses everything claimed, as applied above (see claim 23). Chen further discloses to apply the BIC criterion in clustering (page 8, paragraph 2). Moreover, Chen discloses that in the hierarchical clustering two nodes can be merged only if the merging increases the BIC value (abstract, also see page 9, paragraph 3) that suggests the two models used in identifying step are also applied in clustering step, which is read on the claimed "clustering step is performed using a BIC model-selection criterion, where a first model assumes that two segments or clusters should be merged, and a second model assumes that said two segments or clusters should be maintained independently."

Regarding **claim 29**, Chen discloses everything claimed, as applied above (see claim 23). Chen further discloses that using M segments and k clusters (page 8, paragraphs 2 and 3) successively merge two nearest nodes in clustering step and generate a new cluster set S' from

pervious set S (page 9, paragraph 3), which is read on the claimed "clustering step is performed using K previously identified clusters and M segments to be clustered."

Regarding **claim 30**, it discloses an apparatus, which corresponds to the method of claim 1; the apparatus is inherent in that it simply provides structure for the functionality found in claim 1.

Regarding **claim 31**, it discloses an article of manufacture, which corresponds to the method of claim 1; the article of manufacture is inherent in that it simply provides structure and implementation for the functionality found in claim 1.

Regarding **claim 32**, it discloses an apparatus, which corresponds to the method of claim 16; the apparatus is inherent in that it simply provides structure for the functionality found in claim 16.

Regarding **claim 33**, it discloses an article of manufacture, which corresponds to the method of claim 16; the article of manufacture is inherent in that it simply provides structure and implementation for the functionality found in claim 16.

Regarding **claim 34**, it discloses an apparatus, which corresponds to the method of claim 23; the apparatus is inherent in that it simply provides structure for the functionality found in claim 23.

Regarding **claim 35**, it discloses an article of manufacture, which corresponds to the method of claim 23; the article of manufacture is inherent in that it simply provides structure and implementation for the functionality found in claim 23.

Claim Rejections - 35 USC § 103

The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

2. Claims 6-7, 9, 20 and 27 are rejected under 35 U.S.C. 103(a) as being unpatentable over Chen in view of well known prior art (MPEP 2144.03).

Regarding **claim 6**, Chen discloses everything claimed, as applied above (see claim 5). Chen further cites that by expanding the window, the final decision of a change point is made based on as much data points as possible (page 6, paragraph 2). But, Chen fails to specifically disclose to increase small window size in slow manner and increase larger window size in a faster manner. However, the examiner takes official notice of the fact that it was well known in the art to adjust increase rate based on data size processed.

Therefore, it would have been obvious to one of ordinary skill in the art at time the invention was made to modify Chen by specifically providing an adjustable increase rate base on processed window size, for the purpose of reducing processing time.

Regarding **claim 7**, Chen and well-known prior art disclose everything claimed, as applied above (see claim 6). Chen further discloses that the window size [a=t+1, b=a+1]=1 is reinitialized after detecting a segment boundary (page 6, paragraph 1), which is read on the claimed "window size, n, is initialized to a minimum value after a segment boundary is detected."

Regarding **claim 9**, Chen discloses everything claimed, as applied above (see claim 2). But, Chen fails to specifically disclose that "BIC model selection test is not performed when the window size, n, exceeds a predefined threshold." However, the examiner takes official notice of the fact that it was well known in the art to stop a process when it exceeds a predefined threshold.

Therefore, it would have been obvious to one of ordinary skill in the art at time the invention was made to modify Chen by specifically providing a predefined threshold and a test condition for the purpose of preventing a process from over sizing.

Regarding **claim 20**, Chen discloses everything claimed, as applied above (see claim 17). But, Chen fails to specifically disclose that "wherein said BIC model selection test is not performed where the detection of a boundary is unlikely to occur." However, the examiner takes official notice of the fact that it was well known in the art to skip certain portion of data for processing, because the portion has very small chance to be hit.

Therefore, it would have been obvious to one of ordinary skill in the art at time the invention was made to modify Chen by specifically providing a skipping mechanism for those data that unlikely have a boundary for detection, for the purpose of increasing efficiency and reducing processing time.

Regarding **claim 27**, Chen discloses everything claimed, as applied above (see claim 26). But, Chen fails to specifically disclose that "wherein said BIC model selection test is not performed where the detection of a boundary is unlikely to occur." However, the examiner takes official notice of the fact that it was well known in the art to skip certain portion of data for processing, because the portion has very small chance to be hit.

Therefore, it would have been obvious to one of ordinary skill in the art at time the invention was made to modify Chen by specifically providing a skipping mechanism for those data that unlikely have a boundary for detection, for the purpose of increasing efficiency and reducing processing time.

3. Claim 15 is rejected under 35 U.S.C. 103(a) as being unpatentable over Chen in view of Kleider et al. (USPN 5,930,748), hereinafter referenced as Kleider.

Regarding **claim 15**, Chen discloses everything claimed, as applied above (see claim 1). But, Chen fails to specifically disclose "processing said audio source with a speaker identification engine to assign a speaker name to each of said cluster." However, the examiner contends that the concept of providing an identified speaker cluster with a speaker name was well known, as taught by Kleider.

In the same field of endeavor, Kleider discloses a speaker identification system and method. Kleider employs a speaker identification metric (226) (Fig. 2) in that each element is associated with one particular speaker in the speaker model data 213 (Fig. 2) (column 6, lines 25-32). Kleider further suggests that the information of the speaker model data may include speaker name (column 6, line 44).

Therefore, it would have been obvious to one of ordinary skill in the art at time the invention was made to modify Chen by specifically providing a speaker identification mechanism to associate a speaker cluster identifier with a speaker name, as taught by Kleider, for the purpose of using a common identifier in a speaker identification system.

(11) Response to Argument

Applicant's main arguments (appeal brief: pages 3-4) regards independent claim 1 (similar to claims 16, 23 and 30-35), which recites "a method of tracking a speaker in an audio source, said method comprising the steps of: identifying potential segment boundaries in said audio source; and clustering homogeneous segments from said audio source substantially concurrently with said identifying step" (claim 1).

In response to applicant's argument that the prior art (Chen) discloses "the audio stream is first segmented and then clustered" (appeal brief: page 3, lines 14-15), but "no indication of or suggestion to perform segmentation and clustering 'substantially concurrently' in the cited text" (appeal brief: page 4, lines 14-15), thus "does not disclose or suggest" the claim limitation (recited above) (appeal brief: page 4, lines 9-10), examiner respectfully disagrees with applicant, and has a different view of the prior art teachings and the interpretation of the claimed limitations. It is noted that the claimed limitation comprises two elements (steps), segmentation and clustering, which is disclosed by Chen (Chen: page 1, paragraph 2), as stated in the claim rejection in the final office action; the only argument left is the limitation of "substantially concurrently". First, examiner believes that the limitation "substantially concurrently" has no patentable weight, because applicant does not have any clear definition and/or description in the claim or in the specification about this limitation, and does not give any conditions to apply this limitation. Second, examiner believes that the prior art (Chen) explicitly and/or implicitly discloses all the limitations regarding claim 1, including the limitation of "substantially concurrently", based the interpretation of the claim language and the understanding prior art teachings. It is noted that the limitation of "substantially concurrently" for performing the two

steps may associate many time related factors, such as, computing speed, simple rate, a number of speakers, minimum recognizable data samples, frequency of changing speakers, total stream size, but applicant does not disclose these factors, so that it can be very broadly interpreted. For example, based on the claim, a assumption may includes that: (1) before clustering, certain amount of segments must be segmented, in other words, for identifying speakers, data segmentation is always first preformed, then data clustering is preformed; (2) in order to identify speakers, processed data stream must have at least two data groups that correspond to at least two speakers, which satisfies the claimed limitation and supported by the specification (see specification: page 8, paragraphs 3-4 and page 13, paragraphs 3-4, and Fig.1 and Fig. 2). It is also noted that Chen's disclosure satisfies at least a minimum condition of above assumption, which is only two speakers and two data groups in a speech data stream, because Chen recites "comparing two models, one models the data as two Gaussian(s); the other models the data as just one Gaussian" to detect the changing point for segmentation (Chen, Section 3.1, page 4), which suggests that the segmentation can be used for at least two speakers; similar suggestion is also applied to clustering (Chen, Section 4., page 8). Therefore, at least under this minimum condition, the prior art has the same situation as the applicant claimed, regardless of whatever the limitation "substantially concurrently" exactly means, thus, the prior art does disclose and/or suggest the limitation in claim 1, and it also satisfies the claimed "clustering segments from said audio source corresponding to the same speaker during said pass through said audio source" (claim 23). Therefore, the claim rejection in final office action is proper.

In response to applicant's argument that "the clustering in Chen is performed only after the audio stream has been segmented" and "each segment is compared to all other segments

before clustering is finalized" (appeal brief: page 3, lines 16-18), examiner notes that the claim 1 does not recite these limitation(s), so that they are irrelevant to the claim.

In response to applicant's argument regarding claim 15 under 35 USC 103(a) that the "Kleider et al. do not address the issue of segmenting speech" and "thus, Kleider do not disclose or suggest" the limitation of claim 1 (appeal brief: page 5, paragraphs 2-3), examiner respectfully disagrees with applicant. It is noted that this rejection under the combination of prior arts (Chen and Kleider), one cannot show nonobviousness by attacking references individually where the rejections are based on combinations of references. See *In re Keller*, 642 F.2d 413, 208 USPQ 871 (CCPA 1981); *In re Merck & Co.*, 800 F.2d 1091, 231 USPQ 375 (Fed. Cir. 1986). (also see detail in the claim rejection stated above or in final office action).

For the above reasons, it is believed that the rejections should be sustained.

Respectfully submitted,

December 4, 2003

Conferees

Richemond Dorvil, Vijay Chawan, Qi Han

KEVIN M MASON
RYAN, MASON & LEWIS, LLP
1300 POST ROAD, SUITE 205
FAIRFIELD, CT 06430

RICHEMOND DORVIL
SUPERVISORY PATENT EXAMINER